

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ :	A2	(11) International Publication Number: WO 99/67291
C07K 14/715, A61K 47/48, C07K 1/107		(43) International Publication Date: 29 December 1999 (29.12.99)

(21) International Application Number: PCT/US99/13953	(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 18 June 1999 (18.06.99)	
(30) Priority Data: 09/102,530 22 June 1998 (22.06.98) US	
(71) Applicant (for all designated States except US): IMMUNEX CORPORATION [US/US]; 51 University Street, Seattle, WA 98101 (US).	
(72) Inventor; and	
(73) Inventor/Applicant (for US only): PETTIT, Dean, K. [US/US]; 3844 48th Avenue N.E., Seattle, WA 98105 (US).	
(74) Agent: HENRY, Janis, C.; Immunex Corporation, Law Dept., 51 University Street, Seattle, WA 98101 (US).	

(54) Title: SITE SPECIFIC PROTEIN MODIFICATION BY MUTAGENESIS

1 11 21 31 41
LPAQVAFTPY APEPGSTCRL REYYDQTAQM CCEKCSPOQH AIVFCTKTSQ
51 61 71 81 91
TVCDSCFESI VNGGKWWVDE ECGCGRAGS DQVETQACTR EQNRICTCRP
101 111 121 131 141
G W Y C A L S K Q E T C E L C A P L R K C R P G F G V A R P Q T E T S D V V C K P C A P Q T F S N T
151 161 171 181 191
T S S T D I C R P H Q I C N V V A I P G N A S M D A V C T S T S P T R S M A P G A V H L P V S T R S
201 211 221 231
Q H T Q P T P E P S T A P S T S F L L P M G P S P P A E G S T Q D

(57) Abstract

Processes for conjugating proteins with polyethylene glycol are disclosed. The disclosed processes provide modified proteins having little or no decrease in their activity and include the steps of deleting at least one amino acid residue on the protein, replacing the at least one amino acid residue with an amino acid residue that does not react with polyethylene glycol, and contacting the protein with polyethylene glycol under conditions sufficient to conjugate the polyethylene glycol to the protein. This advantageous retention of a desired protein activity is attributed to the availability of one or more protein binding sites which is unaltered in the conjugation process and thus remains free to interact with a binding partner ligand or cognate subsequent to the conjugation process.

TITLE

SITE SPECIFIC PROTEIN MODIFICATION BY MUTAGENESIS

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to processes for modifying proteins. More particularly, the present invention involves processes for linking polyethylene glycol to proteins in a manner which provides advantages associated with polyethylene glycol conjugated proteins while maintaining a desired protein bioactivity.

Description of Related Art

Processes and reagents for chemically modifying proteins have been used extensively for decades. Traditionally, protein chemical modifications were carried out in order to study their functional properties and structural characteristics. With the emergence of recombinant DNA techniques and interest in protein therapeutics, researchers have chemically modified proteins to improve their clinical performance. In particular, processes for conjugating proteins with polyethylene glycol have gained widespread use within the pharmaceutical and biochemical communities as a result of numerous improved pharmacological and biological properties associated with polyethylene glycol conjugated proteins. For example, polyethylene glycol conjugated proteins are known to have significantly enhanced plasma half life, and thus have substantially improved the clinical usefulness. Additionally, polyethylene glycol conjugated proteins generally have reduced antigenicity and immunogenicity, thereby are less prone to causing life-threatening anaphylaxis.

Another benefit associated with polyethylene glycol conjugated proteins is that of water solubility which is increased as a result of the high water solubility of polyethylene glycol. The increased water solubility can improve the protein's formulation characteristics at physiological pH's and can decrease complications associated with aggregation of low solubility proteins.

Additionally, polyethylene glycol conjugated proteins have found use in bioindustrial applications such as enzyme based reactions in which the reaction environment is not optimal for the enzyme's activity. For example, some polyethylene

SUMMARY OF THE INVENTION

The present invention provides protein modification processes that result in modified proteins having little or no decrease in an activity associated with the protein.

5 More particularly, the invention described herein includes processes for modifying a protein by first deleting one or more amino acid sites on the protein that is suitable for polyethylene glycol conjugation and then contacting the protein with polyethylene glycol under conditions suitable for conjugating the polyethylene glycol to the protein.

10 Preferably, the deleted amino acid residue is replaced with an amino acid residue that does not react with polyethylene glycol. The resulting polyethylene glycol conjugated protein has improved characteristics over proteins modified according to prior art procedures. An advantageous retention of activity is attributed to the availability of one or more protein binding sites which is unaltered in the protein modification process and thus remains free to interact with a binding partner subsequent to the modification

15 process. Within the present invention are proteins useful for polyethylene glycol conjugation and polyethylene glycol conjugated proteins prepared by processes described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

20 Fig. 1 illustrates lysine residues within p75 TNF receptor extracellular domain that are polyethylene glycol conjugation sites and lysine residues that make contact with TNF α .

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides processes and reagents for conjugating proteins or 25 polypeptides with polyethylene glycol in a manner that results in polyethylene glycol conjugated proteins having little or no reduction in a desired activity. More specifically, the present invention provides processes for conjugating polyethylene glycol with proteins under conditions which preclude polyethylene glycol conjugation at one or more selected sites on the protein. When the one or more selected sites is active in a protein 30 binding domain, preventing polyethylene glycol conjugation at the site in accordance with the present invention contributes to maintaining a desired bioactivity while providing benefits associated with polyethylene glycol conjugation.

polyethylene glycol under the reaction conditions. Additionally, the deleted amino acid residue or residues is replaced with an amino acid residue that does not significantly diminish the activity of the native protein. The resulting polyethylene glycol conjugated protein will have an activity associated with selected binding sites and, depending upon 5 the degree to which additional sites are involved in the conjugation process, will have a diminished, or no activity, associated with such additional sites. This approach is useful in cases in which cognate or substrate binding to one or more protein binding sites is desirably suppressed in certain clinical, diagnostic or industrial applications.

Proteins that may be modified in accordance with the present invention include 10 those having utility in clinical and diagnostics applications and those used in the biotechnology industry, such as enzymes in bioreactors. Receptors which may be modified as taught herein include cytokine receptors, for example, TNFR, IL-4R, IL-1R, IL-17R, IL-15R, p55 TNFR:Fc and p75 TNFR:Fc. Candidate antibodies for conjugation include but are not limited to OKT3 (anti-T-Cell), AVAKINE™ (anti-TNF) and anti 15 Her2/Neu. Enzymes of interest for conjugation include CD39, tPA, and DNase. Many proteins have multimeric binding sites and require more than one association for activity. Such proteins are particularly desirable for modification since loss of one binding site leaves the whole protein inactive. Members of the group of multimeric proteins include TNF, hGH, CD40L, and FasL. Other candidate protein ligands are known to bind 20 multiple receptor subunits and include IL-2, IL-15, GM-CSF, and G-CSF.

In accordance with the present invention, a selected amino acid residue is one that is associated with a site on a protein which contributes to a specific function of that protein, and which is reactive with polyethylene glycol under the protein modification reaction conditions. The selected amino acid residue may be directly involved with a 25 binding association with a protein binding partner. Alternatively, the amino acid may be sufficiently central to the spatial configuration of the protein that modifying the protein with polyethylene glycol results in a significant loss of desirable properties even though the amino acid residue is neither within a binding site nor directly or physically involved with the protein's interaction with a binding partner. Sites include, but are not limited to, 30 cognate sites or substrate binding sites which are associated with a protein activity.

Amino acid residues that are reactive with polyethylene glycol under conditions known in the art include those having residues having nucleophilic moieties that are available for reaction with polyethylene glycol or an activated polyethylene glycol. For

After expressing, collecting and purifying the engineered proteins encoded by the mutated DNA, the expressed proteins can be reacted with polyethylene glycol to form a conjugated protein. Then the conjugated protein can be tested for functional activity and other characteristics such as immunogenicity, physiological clearance, and solubility.

5 The polyethylene glycol conjugated proteins that have the desired activity and most favorable clearance, solubility and immunogenicity properties also contain the desired selected lysine residues i.e., the residues that had been deleted and replaced prior to reacting the protein with polyethylene glycol.

For many proteins, the location of polyethylene glycol reactive amino acid residues and their conformational contribution to the structure and function of the protein are known. Among these proteins, are those for which the crystalline structure of the protein is known, and, in some cases, the crystalline structure of the protein-binding partner complex is known. For these proteins, determining a selected amino acid residue typically requires only identifying the residues that are within the protein's binding domain or in close spatial proximity to the protein's binding region and identifying those residues that are reactive with polyethylene glycol under the contemplated polyethylene glycol reaction conditions.

In accordance with the present invention, deleting a selected amino acid residue on the protein can be accomplished with a variety of suitable procedures that provide modified proteins. In the context of the present invention, such procedures include, but are not limited to, site directed mutagenesis techniques and direct protein synthesis methods in which the protein lacking one or more selected amino acid residues is synthesized using standard protein synthesis procedures known in the art. As noted above, preferably the process of deleting a selected amino acid residue additionally involves replacing the selected amino acid residue with an amino acid residue that is not reactive with polyethylene glycol.

Proteins may be prepared by any of a number of conventional techniques. A desired DNA sequence may be chemically synthesized using techniques known per se. DNA fragments also may be produced by restriction endonuclease digestion of a full length cloned DNA sequence, and isolated by electrophoresis on agarose gels. Linkers containing restriction endonuclease cleavage site(s) may be employed to insert the desired DNA fragment into an expression vector, or the fragment may be digested at cleavage sites naturally present therein.

bridge are also potential reaction sites on the protein. The general principles of polyethylene glycol conjugation with proteins, and common activating reagents are described by Delgado et al. in *The Uses and Properties of PEG-Linked Proteins*, from *Critical Reviews in Therapeutic Drug Carrier Systems*, 9(3,4):249-304 (1992) and the 5 ACS Symposium Series 680 ed. y Harris et al., *Poly(ethylene glycol) Chemistry and Biological Applications* 1997, both of which are incorporated herein by reference.

Activated forms of polyethylene glycol and monomethoxypolyethylene glycol are commercially available and may be used in processes of the present invention. Most notably, Shearwater Polymers, Inc of Huntsville, AL provides a number of polyethylene 10 glycol polymers and polyethylene glycol derivatives. The Shearwater Polymers, Inc Catalog (Shearwater Polymers, Inc. Catalog Functionalized Biocompatible Polymers for Research, 1997-1998 incorporated herein by reference) describes and makes available a wide variety of activated polyethylene glycols suitable for coupling with proteins under a wide range of reaction conditions. This catalog additionally provides preferred reaction 15 conditions for their derivatized polyethylene glycol reagents. Those skilled in the art having been made aware of the numerous reagents suitable for conjugating proteins with polyethylene glycol will appreciate the variety of reagent choices in view of the nature of the protein selected, the nature of the reactive amino groups or sulphydryl groups on the protein and the end use of the conjugated protein. For example, to provide conjugated 20 proteins having improved solubility, activity characteristics and delivery properties but not necessarily increased clinical clearance time, a succinimidyl succinate activated polyethylene glycol (SS-PEG) can be used in the conjugation reaction. The ester link to the protein is less stable and will hydrolyze *in vivo*, releasing the polyethylene glycol from the protein. Activated polyethylene glycols are available which will more 25 preferentially react with amino groups as opposed to sulphydryl groups and vice versa. Commonly selected activated polyethylene glycols include succinimidyl carbonate activated polyethylene glycols, succinimidyl succinate activated polyethylene glycol and succinimidyl propionic acid polyethylene glycols.

As an alternative to selecting commercially available activated polyethylene 30 glycols, a polyethylene glycol of interest may be activated using reagents which react with hydroxyl functionalities to form a site reactive with a site on a protein of interest. Typically, the protein reactive site is an amino group but can be a sulphydryl or hydroxyl and the activated polyethylene glycol typically is an active ester or imidazole (See pgs 274 - 285 *ibid.*) Preferably, only one hydroxyl functionality of the polyethylene glycol is

conjugated, their molecular size causes them to extend their spacial or steric influence so that binding or receptor sites have reduced accessibility. It is within the knowledge of those skilled in the art to determine an optimum polyethylene glycol molecular weight for any selected protein and benefits desired from the polyethylene glycol conjugation.

5 While the above described polyethylene glycol conjugation procedures are those in which the result is polyethylene glycol conjugated to protein via a covalent bond, it is within the scope of the present invention to include procedures in which the conjugation is via a different association. In the context of the present invention, proteins may be modified by conjugating them to polyethylene glycol using a variety of different linking 10 or conjugating mechanisms. For example, a protein selected for conjugation can be derivatized at an amino group or other suitably reactive functionality with a polyA oligonucleotide and then conjugated with a polyethylene glycol derivatized with a polyT oligonucleotide. Another approach involves derivatizing the protein with a functionality having a known specific binding partner and then conjugating the protein with 15 polyethylene glycol which has been derivatized with the binding partner for the functionality. For example, a protein can be derivatized with biotin and the polyethylene glycol derivatized with streptavidin or avidin (or vice versa). This results in the specific binding of polyethylene glycol to those protein sites having the biotin. A number of reagents for modifying proteins for the purpose of introducing certain functionalities are 20 commercially available. For example, the Pierce ImmunoTechnology catalogue identifies and provides access to a variety of reagents associated with protein modification. Among these are Traut's Reagents and SATA (Pierce ImmunoTechnology Catalogue, Vol I, pg E-14) which can introduce active groups at N-terminal amines and lysine amino 25 functionalities. These active groups provide sites for further introducing functionalities for reacting more specifically with polyethylene glycol. Those skilled in the art will also recognize that ionic interactions between polyethylene glycol and a protein of interest are also possible. For example, an association between an ionic moiety on the protein and its counter ion on polyethylene glycol can be utilized if the association is sufficiently strong to remain associated under physiological conditions.

30 Further embodiments of the present invention which may utilize prior modified proteins include those processes in which the protein selected for conjugation has too few potential polyethylene glycol conjugation sites or no potential polyethylene glycol conjugation sites outside the protected amino acid region. By modifying the selected protein to introduce amino and sulphydryl sites on the protein sufficient polyethylene

fractions. Because the polyethylene glycol reaction can be taken to completion and all the available polyethylene glycol sites can be fully reacted, the final product is more homogeneous than prior art products which are prepared under conditions that favor reaction at specific sites.

5

The following examples are presented in order to provide a more detailed description of specific embodiments of the present invention and are not to be construed as limiting the scope of the invention.

10

EXAMPLE 1

Selecting a Protein Modification Site

The following describes a procedure for identifying amino acid residues of p75 TNF receptor for deletion and substitution in accordance with the present invention. Because the expected polyethylene glycol modification reaction conditions were to be those that favor modification of the ϵ -amino group of lysine residues and the N-terminal amine, the amino acids identified were lysine residues that make contact between the TNF receptor and the ligand in the TNF receptor-ligand complex.

15

The p75 TNF receptor is from a family of structurally homologous receptors which includes the p55 TNF receptor. TNF α and TNF β (TNF ligands) compete for binding to the p55 and p75 TNF receptors. The x-ray crystal structure of the complex formed by the extracellular domain of the human p55 TNF receptor and TNF β has been determined (Banner et al. *Cell* 73:431, 1993, incorporated herein by reference). This crystallography work confirmed that the complex of p55 TNF receptor and TNF β has three p55 TNF receptor molecules bound symmetrically to one TNF β trimer. The studies further demonstrated that the receptor binds in a groove between two adjacent TNF β subunits. Advantageously, the crystal structure of the complex provides a model for TNF receptor structure and activation and can be used to identify amino acid domains within the ligand and in the receptor that make contact to for the complex.

20

A sequence alignment of the p55 TNF receptor amino acid sequence and the p75 TNF receptor amino acid sequence reveals that p75 TNF receptor residues K34, K42, K47, K108, K120, and K140 are closely aligned with p55 TNF receptor residues K32, Y40, G45, S108, L119 and T138. (See Banner et al. *Cell* 73:431, 1993). Based upon this alignment information and molecular modeling that illustrates the spatial positions of

the nucleotide sequence for the K108R, K120R mutant and SEQ ID NO:6 describes the amino acid sequences encoded by SEQ ID NO:5.

Briefly, the mutants were prepared using site directed mutagenesis of K108 and/or K120 in the human p75 TNF receptor using PCR mutagenesis of the Srf1-Not1 fragment of hTNF receptor and Fc fusion protein (hTNFR:Fc). The mutant TNF receptor fragments were ligated in frame with a human Fc fragment in the mammalian expression vector sf Haveo409. Several of the prepared clones were sequenced to confirm that the desired nucleic acid changes were incorporated into the mutant nucleotide sequences.

More particularly, PCR mutagenesis was used to generate mutated 430 base pair Sal/Srf1 fragments. The PCR mutagenesis procedures utilized wild type TNFR cDNA (SEQ ID NO:7) used as the template for the PCR reactions. The oligonucleotide sequences used in the PCR reactions to generate the 3 mutant Sal1-Srf1 DNA fragments were as follows:

For the TNF receptor (K108R) mutant the 3' oligonucleotide contained an A to G substitution at position 389 and a Srf1 site at the 3' end. For the TNF receptor (K120R) mutant the 3' oligonucleotide contained an A to G substitution at position 425 and a Srf1 site at the 3' end. For the TNF receptor (K108R,K120R) mutant the oligonucleotide contained an A to G substitution at position 389 and 425 and a Srf1 site at the 3' end. The 5' oligonucleotide used to generate the mutant PCR DNA fragments had no nucleotide changes in the TNFR coding nucleotides and contained the 5' Sal1 site.

For the PCR Reactions the Boehringer Mannheim Expand High Fidelity PCR kit and reagents were used according to manufacturer's directions. The PCR cycling protocol involved the following conditions: 94°C for 2 minutes; 94°C for 30 seconds; 50°C for 15 seconds, 72°C for 1 minute. 25 cycle reaction.

The DNA fragments generated in the PCR reactions were separated on a 1% agarose gel and the 430 base pair TNFR fragments were isolated using GeneClean reagent from BIO101. The isolated fragments were restriction digested with Sal1 and Srf1 from NEB in Universal Restriction Buffer from Stratagene. The DNA was then repurified using the GeneClean reagents from BIO101.

Each of the mutant Sal1/Srf1 DNA 430 fragments generated above (and corresponding to the 5' end of the TNF receptor) was individually ligated with the 1065 basepair Srf1/Not1 DNA fragment corresponding to the 3' TNF receptor and human Fc cDNA and the 7730 basepair Sal1/Not1 pDC409 expression. 20ng of the pDC409 vector

EXAMPLE 3**Conjugating Wildtype and Mutant p75 TNF:Fc Receptors with Polyethylene Glycol**

The following describes a process for preparing polyethylene glycol conjugated wildtype TNFR:Fc molecules and polyethylene glycol conjugated mutant TNFR:Fc molecules. For each polyethylene glycol conjugation reaction, a one hundred micrograms (100 μ g) portion of wildtype TNFR:Fc, or mutant TNFR:Fc, prepared in Example 2 was dissolved in 400 μ L of 50 mM Na₂HPO₄ at pH 8.5 and allowed to react with SPA-PEG 5000 at different molar ratios of polyethylene glycol to protein (calculated as number of lysine residues in TNFR:Fc) overnight at 4°C. The molar ratios of protein to lysine residues 1:1 and 10:1. SPA-PEG is a 5,000 MW succinimidyl carbonate activated monomethoxypolyethylene glycol purchased from Shearwater Polymers, Birmingham, AL. The protein and polyethylene glycol solutions were allowed to react overnight at 2-8°C.

Each of the polyethylene glycol conjugated TNFR:Fc molecules was purified by ion exchange chromatography using SP Sepharose Fast Flow resin (Pharmacia) equilibrated with 20 mM sodium phosphate, pH 7.4. Polyethylene glycol conjugated TNFR:Fc bound to the resin under these conditions. Unreacted polyethylene glycol and reaction byproducts were rinsed from the column with 5 column volumes of the equilibration buffer. The polyethylene glycol conjugated TNFR:Fc was eluted from the column with five column volumes of 20 mM sodium phosphate, 200 mM NaCl, pH 7.4. The eluted fractions were pooled and concentrated to approximately 1-5 mg/mL.

The following indicates the designation given each of the TNFR:Fc molecules conjugated with polyethylene glycol (PEG) by the above described procedure:

- 25 1. PEG-TNFR:Fc(K108R, K120R);
2. PEG-TNFR:Fc(K108R);
3. PEG-TNFR:Fc(K120R);
4. PEG-TNFR:Fc.

30

EXAMPLE 4
Characterization of Conjugated TNFR:Fc

The following describes the characterization of polyethylene glycol conjugated wildtype polyethylene glycol conjugated mutant TNFR:Fc molecules prepared in Example 3 and a control characterization of unconjugated wildtype and mutant TNFR:Fc

Example 5**Pharmacokinetics of Wildtype and Mutant TNFR:Fc Molecules**

The following describes experiments designed to compare the pharmacokinetics of wildtype TNFR:Fc with the polyethylene glycol conjugated TNFR:Fc mutant molecule 5 K108R,K120R (the lysine at 108 and 120 substituted with arginine). The mutant molecule had been conjugated with a polyethylene glycol:lysine ration of 10:1.

Groups of 2 10-12 week old female BALB/c mice were injected intravenously with 10 μ g of wildtype TNFR:Fc or conjugated mutant TNFR:Fc in a total volume of 100 μ L. Following the injection, mice were sacrificed and blood samples were collected 10 at 5 minutes, 1 hour, 8 hours, 24 hours, 48 hours and 72 hours via cardiac puncture. Plasma samples were analyzed by A375 bioassay. The elimination half lives, $t_{1/2}$, of the polyethylene conjugated mutant and the wildtype TNFR:Fc were determined. The half-life values are presented as $t_{1/2}$ +/- S.E. where S.E. indicates the standard error in fitting the 15 log linear line to the data points. The $t_{1/2}$ of wildtype TNFR:Fc was determined to be 16.5 +/- 1.0 hours and that of the polyethylene glycol mutant was determined to be 36.5 +/- 8.5 hours.

The results of the above experiments demonstrate that polyethylene glycol conjugated TNF receptor prepared in accordance with the present invention has a significantly enhanced circulation half life compared to a TNF receptor that is not 20 polyethylene glycol conjugated.

Example 6**Bioactivity of Polyethylene Conjugated Wild type TNFR:Fc and Polyethylene Conjugated Mutant TNFR:Fc**

25

The bioactivities of the polyethylene glycol conjugated TNFR:Fc molecules prepared in Example 3 were measured by *in vitro* A375 bioassays. This assay is generally described in Onozaki et al. *J. Immunology* 135:3962 (1985) and Nakai et al. *Biochem. Biophys. Res. Comm.* 154:1189 (1988). The bioassay is based upon the inhibitory 30 response of the A375 human malignant melanoma adherent cell line to TNF α . Soluble TNFR:Fc can specifically neutralize the inhibitory activity of TNF α in a dose dependent manner. To perform the bioassay, 375 cell line (ATCC CRL 1872) was harvested using a trypsin-EDTA solution to remove the cell monolayer from flasks. The harvested cells were washed with an assay medium of Dulbeccos' Modified Eagles Medium with added

What is claimed is:

1. A process for conjugating a protein with polyethylene glycol, said process comprising the steps of
 - 5 deleting at least one amino acid residue on the protein, and
 - contacting the protein with polyethylene glycol under conditions sufficient to conjugate the polyethylene glycol to the protein.
- 10 2. The process of claim 1 further including the step of replacing the at least one amino acid residue with an amino acid residue that does not react with polyethylene glycol.
- 15 3. The process of claim 2 wherein the one or more amino acid residues is associated with protein binding or protein association.
4. The process of claim 2 wherein the at least one amino acid residue is lysine.
5. The process of claim 2 wherein the amino acid residue that does not react with
- 20 polyethylene glycol is arginine.
6. The process of claim 1 wherein the step of contacting the protected protein with polyethylene glycol comprises causing an activated polyethylene glycol to react with nucleophiles on the protein.
- 25 7. A polyethylene glycol conjugated protein prepared according to the process of claim 1.
8. A process for modifying a protein, said process comprising the steps of
 - 30 deleting at least one amino acid residue on the protein;
 - replacing the at least one amino acid residue with an amino acid residue that does not react with polyethylene glycol; and
 - contacting the protein with polyethylene glycol under conditions sufficient to conjugate the polyethylene glycol to the protein.

17. A DNA encoding mutant soluble TNFR polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4 and SEQ ID NO:6.
- 5 18. A mutant polypeptide of claim 16 wherein the polypeptide is polyethylene glycol conjugated.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Pettit, Dean
- (ii) TITLE OF INVENTION: Site Specific Protein Modification
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Janis C Henry
 - (B) STREET: 51 University
 - (C) CITY: Seattle
 - (D) STATE: WA
 - (E) COUNTRY: US
 - (F) ZIP: 98101
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE: 18 June 1999
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Henry, Janis C
 - (B) REGISTRATION NUMBER: 34,347
 - (C) REFERENCE/DOCKET NUMBER: 2637-WO
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (206)470-4189

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 705 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (v) FRAGMENT TYPE: N-terminal
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..705

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 235 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Leu Pro Ala Gln Val Ala Phe Thr Pro Tyr Ala Pro Glu Pro Gly Ser
 1 5 10 15

Thr Cys Arg Leu Arg Glu Tyr Tyr Asp Gln Thr Ala Gln Met Cys Cys
 20 25 30

Ser Lys Cys Ser Pro Gly Gln His Ala Lys Val Phe Cys Thr Lys Thr
 35 40 45

Ser Asp Thr Val Cys Asp Ser Cys Glu Asp Ser Thr Tyr Thr Gln Leu
 50 55 60

Trp Asn Trp Val Pro Glu Cys Leu Ser Cys Gly Ser Arg Cys Ser Ser
 65 70 75 80

Asp Gln Val Glu Thr Gln Ala Cys Thr Arg Glu Gln Asn Arg Ile Cys
 85 90 95

Thr Cys Arg Pro Gly Trp Tyr Cys Ala Leu Ser Arg Gln Glu Gly Cys
 100 105 110

Arg Leu Cys Ala Pro Leu Arg Lys Cys Arg Pro Gly Phe Gly Val Ala
 115 120 125

Arg Pro Gly Thr Glu Thr Ser Asp Val Val Cys Lys Pro Cys Ala Pro
 130 135 140

Gly Thr Phe Ser Asn Thr Thr Ser Ser Thr Asp Ile Cys Arg Pro His
 145 150 155 160

Gln Ile Cys Asn Val Val Ala Ile Pro Gly Asn Ala Ser Met Asp Ala
 165 170 175

Val Cys Thr Ser Thr Ser Pro Thr Arg Ser Met Ala Pro Gly Ala Val
 180 185 190

His Leu Pro Gln Pro Val Ser Thr Arg Ser Gln His Thr Gln Pro Thr
 195 200 205

Pro Glu Pro Ser Thr Ala Pro Ser Thr Ser Phe Leu Leu Pro Met Gly
 210 215 220

Pro Ser Pro Pro Ala Glu Gly Ser Thr Gly Asp
 225 230 235

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 705 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

CCA GAA CCC AGC ACT GCT CCA AGC ACC TCC TTC CTG CTC CCA ATG GGC	672	
Pro Glu Pro Ser Thr Ala Pro Ser Thr Ser Phe Leu,Leu Pro Met Gly		
210	220	
CCC AGC CCC CCA GCT GAA GGG AGC ACT GGC GAC	705	
Pro Ser Pro Pro Ala Glu Gly Ser Thr Gly Asp		
225	230	235

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 235 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu Pro Ala Gln Val Ala Phe Thr Pro Tyr Ala Pro Glu Pro Gly Ser			
1	5	10	15
Thr Cys Arg Leu Arg Glu Tyr Tyr Asp Gln Thr Ala Gln Met Cys Cys			
20	25	30	
Ser Lys Cys Ser Pro Gly Gln His Ala Lys Val Phe Cys Thr Lys Thr			
35	40	45	
Ser Asp Thr Val Cys Asp Ser Cys Glu Asp Ser Thr Tyr Thr Gln Leu			
50	55	60	
Trp Asn Trp Val Pro Glu Cys Leu Ser Cys Gly Ser Arg Cys Ser Ser			
65	70	75	80
Asp Gln Val Glu Thr Gln Ala Cys Thr Arg Glu Gln Asn Arg Ile Cys			
85	90	95	
Thr Cys Arg Pro Gly Trp Tyr Cys Ala Leu Ser Lys Gln Glu Gly Cys			
100	105	110	
Arg Leu Cys Ala Pro Leu Arg Arg Cys Arg Pro Gly Phe Gly Val Ala			
115	120	125	
Arg Pro Gly Thr Glu Thr Ser Asp Val Val Cys Lys Pro Cys Ala Pro			
130	135	140	
Gly Thr Phe Ser Asn Thr Thr Ser Ser Thr Asp Ile Cys Arg Pro His			
145	150	155	160
Gln Ile Cys Asn Val Val Ala Ile Pro Gly Asn Ala Ser Met Asp Ala			
165	170	175	
Val Cys Thr Ser Thr Ser Pro Thr Arg Ser Met Ala Pro Gly Ala Val			
180	185	190	
His Leu Pro Gln Pro Val Ser Thr Arg Ser Gln His Thr Gln Pro Thr			
195	200	205	
Pro Glu Pro Ser Thr Ala Pro Ser Thr Ser Phe Leu Leu Pro Met Gly			
210	215	220	
Pro Ser Pro Pro Ala Glu Gly Ser Thr Gly Asp			
225	230	235	

GTC TGC ACG TCC ACG TCC CCC ACC CGG AGT ATG GCC CCA GGG GCA GTA	576
Val Cys Thr Ser Thr Ser Pro Thr Arg Ser Met Ala Pro Gly Ala Val	
180	185
190	
CAC TTA CCC CAG CCA GTG TCC ACA CGA TCC CAA CAC ACG CAG CCA ACT	624
His Leu Pro Gln Pro Val Ser Thr Arg Ser Gln His Thr Gln Pro Thr	
195	200
205	
CCA GAA CCC AGC ACT GCT CCA AGC ACC TCC TTC CTG CTC CCA ATG GGC	672
Pro Glu Pro Ser Thr Ala Pro Ser Thr Ser Phe Leu Leu Pro Met Gly	
210	215
220	
CCC AGC CCC CCA GCT GAA GGG AGC ACT GGC GAC	705
Pro Ser Pro Pro Ala Glu Gly Ser Thr Gly Asp	
225	230
235	

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 235 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Leu Pro Ala Gln Val Ala Phe Thr Pro Tyr Ala Pro Glu Pro Gly Ser	
1	5
10	15
Thr Cys Arg Leu Arg Glu Tyr Tyr Asp Gln Thr Ala Gln Met Cys Cys	
20	25
30	
Ser Lys Cys Ser Pro Gly Gln His Ala Lys Val Phe Cys Thr Lys Thr	
35	40
45	
Ser Asp Thr Val Cys Asp Ser Cys Glu Asp Ser Thr Tyr Thr Gln Leu	
50	55
60	
Trp Asn Trp Val Pro Glu Cys Leu Ser Cys Gly Ser Arg Cys Ser Ser	
65	70
75	80
Asp Gln Val Glu Thr Gln Ala Cys Thr Arg Glu Gln Asn Arg Ile Cys	
85	90
95	
Thr Cys Arg Pro Gly Trp Tyr Cys Ala Leu Ser Arg Gln Glu Gly Cys	
100	105
110	
Arg Leu Cys Ala Pro Leu Arg Arg Cys Arg Pro Gly Phe Gly Val Ala	
115	120
125	
Arg Pro Gly Thr Glu Thr Ser Asp Val Val Cys Lys Pro Cys Ala Pro	
130	135
140	
Gly Thr Phe Ser Asn Thr Thr Ser Ser Thr Asp Ile Cys Arg Pro His	
145	150
155	160
Gln Ile Cys Asn Val Val Ala Ile Pro Gly Asn Ala Ser Met Asp Ala	
165	170
175	
Val Cys Thr Ser Thr Ser Pro Thr Arg Ser Met Ala Pro Gly Ala Val	
180	185
190	
His Leu Pro Gln Pro Val Ser Thr Arg Ser Gln His Thr Gln Pro Thr	
195	200
205	

GGG ACG TTC TCC AAC ACG ACT TCA TCC ACG GAT ATT TGC AGG CCC CAC	480
Gly Thr Phe Ser Asn Thr Thr Ser Ser Thr Asp Ile Cys Arg Pro His	
145 150 155 160	
CAG ATC TGT AAC GTG GTG GCC ATC CCT GGG AAT GCA AGC ATG GAT GCA	528
Gln Ile Cys Asn Val Val Ala Ile Pro Gly Asn Ala Ser Met Asp Ala	
165 170 175	
GTC TGC ACG TCC ACG TCC CCC ACC CGG AGT ATG GCC CCA GGG GCA GTA	576
Val Cys Thr Ser Thr Ser Pro Thr Arg Ser Met Ala Pro Gly Ala Val	
180 185 190	
CAC TTA CCC CAG CCA GTG TCC ACA CGA TCC CAA CAC ACG CAG CCA ACT	624
His Leu Pro Gln Pro Val Ser Thr Arg Ser Gln His Thr Gln Pro Thr	
195 200 205	
CCA GAA CCC AGC ACT GCT CCA AGC ACC TCC TTC CTG CTC CCA ATG GGC	672
Pro Glu Pro Ser Thr Ala Pro Ser Thr Ser Phe Leu Leu Pro Met Gly	
210 215 220	
CCC AGC CCC CCA GCT GAA GGG AGC ACT GGC GAC	705
Pro Ser Pro Pro Ala Glu Gly Ser Thr Gly Asp	
225 230 235	

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 235 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Leu Pro Ala Gln Val Ala Phe Thr Pro Tyr Ala Pro Glu Pro Gly Ser	
1 5 10 15	
Thr Cys Arg Leu Arg Glu Tyr Tyr Asp Gln Thr Ala Gln Met Cys Cys	
20 25 30	
Ser Lys Cys Ser Pro Gly Gln His Ala Lys Val Phe Cys Thr Lys Thr	
35 40 45	
Ser Asp Thr Val Cys Asp Ser Cys Glu Asp Ser Thr Tyr Thr Gln Leu	
50 55 60	
Trp Asn Trp Val Pro Glu Cys Leu Ser Cys Gly Ser Arg Cys Ser Ser	
65 70 75 80	
Asp Gln Val Glu Thr Gln Ala Cys Thr Arg Glu Gln Asn Arg Ile Cys	
85 90 95	
Thr Cys Arg Pro Gly Trp Tyr Cys Ala Leu Ser Lys Gln Glu Gly Cys	
100 105 110	
Arg Leu Cys Ala Pro Leu Arg Lys Cys Arg Pro Gly Phe Gly Val Ala	
115 120 125	
Arg Pro Gly Thr Glu Thr Ser Asp Val Val Cys Lys Pro Cys Ala Pro	
130 135 140	
Gly Thr Phe Ser Asn Thr Thr Ser Ser Thr Asp Ile Cys Arg Pro His	
145 150 155 160	